

SUB
D1
CONT.
C1
CONT.

i) encoding two or more related biological molecules into a data structure of initial character strings to provide a collection of two or more different initial character strings wherein each of said biological molecules comprises at least about 10 subunits;
ii) selecting at least two substrings from said initial character strings;
iii) concatenating said substrings to form one or more product strings about the same length as one or more of the initial character strings;
iv) adding the product strings to a data structure to populate a data structure of product strings;
v) optionally repeating steps (i) or (ii) through (iv) using one or more of said product strings as an initial string in the collection of initial character strings, and;
vi) making one or more product biological molecule corresponding to one or more of the product strings.

2 (Amended). The method of claim 1, wherein said encoding comprises encoding two or more nucleic acid sequences into said character strings.

SUB
E3

3 (Amended). The method of claim 2, wherein said two or more nucleic acid sequences comprise a nucleic acid sequence encoding a naturally occurring protein.

4 (Amended). The method of claim 1, wherein said encoding comprises encoding two or more amino acid sequences into said character strings.

5 (Amended). The method of claim 4, wherein said one or more amino acid sequences comprise an amino acid sequence encoding a naturally occurring protein.

SUB
E3
C2

10 (Amended). The method of claim 1, wherein said selecting comprises aligning two or more of said initial character strings to maximize pairwise identity between two or more substrings of the initial character strings, and selecting a character that is a member of an aligned pair for the end of one of the two or more substrings.

11 (Amended). The method of claim 1, wherein said product strings are added to the data structure of product strings only if they have greater than 30% sequence identity with the initial strings.

12 (Amended). The method of claim 1, wherein said method further comprises randomly altering one or more characters of said initial or product character strings.

SUB
E3

14 (Amended). The method of claim 1, wherein said encoding, selecting, or concatenating is performed on an internet site.

15 (Amended). The method of claim 1, wherein said encoding, selecting, or concatenating is performed on a server.

16 (Amended). The method of claim 1, wherein said encoding, selecting, or concatenating is performed on a client linked to a network.

SUB
D2

17 (Amended). A computer program product on a computer readable media comprising computer code that:

SUB
DA
CONT.

i) encodes two or more biological molecules into initial character strings to provide a collection of two or more different initial character strings wherein each of said biological molecules comprises at least about ten subunits;
ii) selects at least two initial substrings from said character strings;
iii) concatenates said substrings to form one or more product strings about the same length as one or more of the initial character strings;
iv) adds the product strings to a data structure to populate a data structure of product strings;
v) optionally repeats steps (i) or (ii) through (iv) using one or more of said product strings as an initial string in the collection of initial character strings; and,
vi) directs the production of one or more product biological molecule corresponding to the product strings.

C2
CONT.

18 (Amended). The computer program product of claim 17, wherein said two or more biological molecules are nucleic acid sequences.

19 (Amended). The computer program product of claim 17, wherein said two or more biological molecules are nucleic acid sequences of naturally occurring proteins.

20 (Amended). The computer program product of claim 17, wherein said two or more biological molecules are amino acid sequences.

SUB
E3

21 (Amended). The computer program product of claim 17, wherein said biological molecules have at least 30% sequence identity.

22 (Amended). The computer program product of claim 17, wherein said computer code selects substrings such that the ends of said substrings occur in string regions of about three to about twenty characters that have higher sequence identity with a corresponding region of another of said initial character strings than the overall sequence identity between the two substrings.

23 (Amended). The computer program product of claim 17, wherein said computer code selects substrings such that the ends of said substrings occur in predefined motifs of about 4 to about 8 characters.

24 (Amended). The computer program product of claim 17, wherein said computer code selects and concatenates substrings from two different initial strings such that the concatenation occurs in a region of about three to about twenty characters having higher sequence identity between said two different initial strings than the overall sequence identity between said two different initial strings.

25 (Amended). The computer program product of claim 17, wherein the computer code selects substrings by aligning two or more of said initial character strings to maximize pairwise identity between two or more substrings of the character strings, and selecting a character that is a member of an aligned pair for the end of one substring.

26 (Amended). The computer program product of claim 17, wherein said product strings are added to the collection only if they have greater than 30% identity with the initial strings.

27 (Amended). The computer program product of claim 17, wherein said computer code additionally randomly alters one or more characters of said character strings.

28 (Amended). The computer program product of claim 27, wherein said computer code additionally randomly selects and alters one or more occurrences of a particular preselected character in said character strings.

29 (Amended). The computer program product of claim 17, wherein said computer code is stored on media selected from the group consisting of magnetic media, optical media, optomagnetic media.

30 (Amended). The computer program product of claim 17, wherein said computer code is in dynamic or static memory of a computer.

NEW CLAIMS

Please add the following new claims:

45 (New). The method of claim 1, wherein the initial character strings of (i) are related.

46 (New). The method of claim 1, further comprising physically screening the molecule(s) represented by the product strings for one or more desired properties.

47 (New). The method of claim 1, further comprising determining a computationally predicted property for molecules represented by the product strings.

48 (New). The method of claim 1, wherein the molecules represented by the product strings are made in parallel in an array of vessels.

49 (New). The method of claim 1, wherein the molecules represented by the product strings are made by assembly of oligonucleotides.

50 (New). The method of claim 1, further comprising testing members of the data structure of product strings for a particular property and determining an optimal combination of sequences using multi-variate analysis.

51 (New). The computer program product of claim 17, wherein the initial character strings of (i) are related.

52 (New). The computer program product of claim 17, wherein the code instructs physical screening of the molecule(s) represented by the product strings for one or more desired properties.

SUB
E3
C3
conc

53 (New). The computer program product of claim 17, wherein the code instructs determination of a computationally predicted property for molecules represented by the product strings.

54 (New). The computer program product of claim 17, wherein the molecules represented by the product strings are made in parallel in an array of vessels.

55 (New). The computer program product of claim 17, wherein the molecules represented by the product strings are made by assembly of oligonucleotides.

56 (New). The computer program product of claim 17, wherein the code tests members of the data structure of product strings for a particular property and determines an optimal combination of initial sequences using multi-variate analysis.

These amendments are made without prejudice and are not to be construed as abandonment of the previously claimed subject matter or agreement with any objection or rejection of record. In accordance with the requirements of 37 C.F.R. § 1.121, a marked up version showing the changes to the claims, is attached herewith as Appendix A. For the Examiner's convenience, a complete claim set of the currently pending claims is also submitted herewith as Appendix B.

REMARKS

Claims 1-30 and 45-56 are pending with entry of this amendment, claims 31-44 being withdrawn pursuant to restriction. The amendments to the claims and the new claims introduce no new matter. Support for the amendments is replete throughout the specification. For example, support for new claims to use of related sequences in the methods is found at page 12, lines 12- 18. Support for introducing variation into initial or product strings is found, e.g., at page 26, lines 28-31. Support for screening is found, e.g., at page 38, lines 17-18 and e.g., at page 42, lines 18-19. Support for testing for properties in silico is found, e.g., at page 42, lines 16-18 and at page 25, lines 26-32. Support for synthesis in arrays of vessels is found, e.g., at page 39, lines 5-9. Support for the use of robotics in the methods is found, e.g., at page 39, lines 12-19. Support for assembly of nucleic acids of interest by oligo assembly is found, e.g., at page 37, lines 5-23. Support for screening for optimal sequences is found, e.g., at page 42, lines 20-31. Support for multi-variate analysis in the methods is found, e.g., at page 42, lines 20-31.